



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/829,831	04/10/2001	Elizabeth Shriberg	SRI/4316	1269

7590 05/25/2005

Thomason, Moser & Patterson LLP  
Attorneys At Law  
First Floor  
595 Shrewsbury Avenue  
Shrewsbury, NJ 07702

EXAMINER
----------

ALBERTALLI, BRIAN LOUIS

ART UNIT	PAPER NUMBER
----------	--------------

2655

DATE MAILED: 05/25/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

**Office Action Summary**

Application No.

09/829,831

Applicant(s)

SHRIBERG ET AL.

Examiner

Brian L Albertalli

Art Unit

2655

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

- 1) ☒ Responsive to communication(s) filed on 12 January 2005.
- 2a) ☒ This action is **FINAL**. 2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

- 4) ☒ Claim(s) 1-21 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1-21 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

**Application Papers**

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on \_\_\_\_\_ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).  
a) ☐ All b) ☐ Some \* c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
- \* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftperson's Patent Drawing Review (PTO-948)
- 3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)  
Paper No(s)/Mail Date \_\_\_\_\_
- 4) ☐ Interview Summary (PTO-413)  
Paper No(s)/Mail Date. \_\_\_\_\_
- 5) ☐ Notice of Informal Patent Application (PTO-152)
- 6) ☐ Other: \_\_\_\_\_

## **DETAILED ACTION**

### ***Response to Amendment***

1. The amendments to the claims have been entered. Claims 19 and 20 are currently amended.

### ***Response to Arguments***

2. Applicant's arguments filed January 12, 2005 have been fully considered but they are not persuasive.

Regarding independent claims 1, 11, and 21, the Applicant argues (see page 8 of Applicant's remarks) that Pickering (U.S. Patent 6,496,799) does not disclose "producing an endpoint signal".

As broadly recited in the claim, "producing an endpoint signal corresponding to the occurrence of the at least one speech endpoint" has been interpreted by the Examiner as any signal that indicates an endpoint has been detected. As highlighted by the Applicant, Pickering discloses testing whether or not a user utterance has been completed using prosodic features (see column 10, lines 21-23 of Pickering). The result of this test must inherently produce a "signal" indicating the occurrence of a speech endpoint. For example, in Fig. 3, step 560, the test to determine whether an endpoint has been detected is performed. As a result of this test, the method either branches to step 570 to perform further actions (Yes branch) or returns to step 520 to receive further caller input (No branch). The output to the "Yes" branch is a "signal" that indicates an endpoint has been detected. Especially when this method is implemented by the

necessary digital circuitry, the "Yes" branch following step 560 produces a "signal" (the binary signal output by the processor that would correspond with the completion of that step of the method). The suggestion by the Applicant that Pickering "merely performs a test" is unpersuasive, because a test without some type of output to indicate the result of the test (i.e. a "signal") is inherently useless.

Furthermore, in response to the argument that Pickering does not teach a *separate* endpoint signal in order to facilitate *subsequent* speech recognition processing, it is noted these features are not recited in the rejected claim(s). Although the claims are interpreted in light of the specification, limitations from the specification are not read into the claims. See *In re Van Geuns*, 988 F.2d 1181, 26 USPQ2d 1057 (Fed. Cir. 1993).

Regarding claims 4-5 and 14-15 the Applicant has argued (see pages 9-11 of Applicant's remarks) that the combination of Pickering and Sonmez et al. (*Modeling Dynamic Prosodic Variation for Speaker Verification*) does not disclose the step of generating an endpoint signal, but as discussed above in reference to claims 1, 11, and 21, this step is disclosed by Pickering, and thus the argument is considered moot.

Further, the Applicant has argued that there is no motivation to combine Pickering and Sonmez, because Pickering teaches a method for identifying the completion of a speech signal and Sonmez teaches a method for identifying the speaker. In response to Applicant's argument that there is no suggestion to combine the references, the examiner recognizes that obviousness can only be established by

Art Unit: 2655

combining or modifying the teachings of the prior art to produce the claimed invention where there is some teaching, suggestion, or motivation to do so found either in the references themselves or in the knowledge generally available to one of ordinary skill in the art. See *In re Fine*, 837 F.2d 1071, 5 USPQ2d 1596 (Fed. Cir. 1988) and *In re Jones*, 958 F.2d 347, 21 USPQ2d 1941 (Fed. Cir. 1992). In this case, Sonmez teaches that extracting the pitch slope provides significant data reduction (see page 2, section 3, lines 4-5 of Sonmez). In the speech processing art, it is well established that significant amounts of data must be processed in order to extract meaningful information from the speech signal to provide adequate results. This is true whether the application is voice recognition (identifying who is speaking), accurately detecting endpoints, or speech recognition (identifying what was said). Therefore, any technique for significantly reducing data would be advantageously employed in any speech processing application.

Regarding claims 6 and 16 the Applicant has argued (see pages 11-13 of Applicant's remarks) that the combination of Pickering, Sonmez et al. and Shriberg et al. (*Prosody-Based Automatic Segmentation of Speech Into Sentences and Topics*) does not disclose the step of generating an endpoint signal, but as discussed above in reference to claims 1, 11, and 21, this step is disclosed by Pickering, and thus the argument is considered moot.

Further, the Applicant has argued that there is no motivation to combine Pickering Sonmez, and Shriberg. In response to Applicant's argument that there is no

Art Unit: 2655

suggestion to combine the references, the examiner recognizes that obviousness can only be established by combining or modifying the teachings of the prior art to produce the claimed invention where there is some teaching, suggestion, or motivation to do so found either in the references themselves or in the knowledge generally available to one of ordinary skill in the art. See *In re Fine*, 837 F.2d 1071, 5 USPQ2d 1596 (Fed. Cir. 1988) and *In re Jones*, 958 F.2d 347, 21 USPQ2d 1941 (Fed. Cir. 1992). In this case, the motivation for combining Pickering and Sonmez has been discussed above in reference to claims 4-5 and 14-15. Regarding the additional combination of Shriberg, Shriberg teaches that the baseline parameter is the most useful parameter when evaluating the pitch features of input speech (see page 135, 1<sup>st</sup> column, 2<sup>nd</sup> paragraph, lines 8-16 of Shriberg). Similarly, Pickering evaluates pitch features of input speech. Therefore, the modifying Pickering to use the baseline parameter would ensure the most accurate evaluation of the pitch features of the input speech, regardless of the application.

Regarding claims 7-9 and 17-19, the Applicant has argued (see pages 13-14 of Applicant's remarks) that Pickering does not disclose the step of "producing an endpoint signal", but as discussed above in reference to claims 1, 11, and 21, this step is disclosed by Pickering, and thus the argument is considered moot.

Regarding claims 10 and 20, the Applicant has argued (see pages 14-15 of Applicant's remarks) that Pickering does not disclose the step of "producing an endpoint

Art Unit: 2655

signal", but as discussed above in reference to claims 1, 11, and 21, this step is disclosed by Pickering, and thus the argument is considered moot.

Further, the applicant has argued that Shin et al. (*Speech/Non-Speech Classification Using Multiple Features for Robust Endpoint Detection*) does not teach that improved endpointing may be addressed by analyzing prosody, thus there is no motivation to combine Pickering and Shin.

However, the step of analyzing prosody to produce an accurate endpoint signal is met by Pickering. Shin et al. teaches that any method of accurate endpoint detection increases the speech recognition performance, therefore, using the prosody based endpoint detection as disclosed by Pickering to perform a speech recognition routine would result in a more accurate speech recognizer, and therefore, would be an obvious modification of Pickering.

3. Therefore, the rejections made in the previous Office Action stand.

***Claim Rejections - 35 USC § 102***

4. The text of those sections of Title 35, U.S. Code not included in this action can be found in a prior Office action.

5. Claims 1-3 and 10-13 are rejected under 35 U.S.C. 102(e) as being anticipated by Pickering (U.S. Patent 6,496,799).

In regard to claims 1, 11, and 21, Pickering discloses a method, apparatus (computer workstation), and electronic storage medium for processing a speech signal comprising:

- extracting prosodic features from a speech signal (spoken pitch);
- modeling the prosodic features to identify at least one speech endpoint (fundamental frequency is derived and then low pass filtered to find gross pitch movements, column 10, lines 30-40); and
- producing an endpoint signal corresponding to the occurrence of the at least one speech endpoint (long decline in pitch value indicates end of the input, column 10, lines 21-23).

In regard to claims 2 and 12, Pickering discloses processing pitch information within the speech signal (column 10, lines 30-40).

In regard to claims 3 and 13, Pickering discloses determining a duration pattern (a test is made to see whether or not the input is silence, column 8, lines 21-22); and performing a pause analysis (system checks whether the amount of silence exceeds a predetermined time-out period, column 8, lines 22-24).



***Claim Rejections - 35 USC § 103***

6. The text of those sections of Title 35, U.S. Code not included in this action can be found in a prior Office action.

7. Claims 4-5 and 14-15 are rejected under 35 U.S.C. 103(a) as being unpatentable over Pickering, in view of Sonmez et al. (*Modeling Dynamic Prosodic Variation for Speaker Verification*).

Pickering is silent as to the details of how the pitch information in the signal is processed.

Sonmez et al. discloses generating a pitch contour (page 2, 1<sup>st</sup> column, second paragraph, third paragraph, and equations 1 and 2);

producing a pitch movement model from the pitch contour; and

extracting a pitch movement slope from the pitch movement model (page 2, section 3, first paragraph and segment slope equation).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Pickering to extract pitch slope from the pitch movement model, since the stylized contours provide significant data reduction, as taught by Sonmez et al. (page 2, section 3, lines 4-5).

Art Unit: 2655

8. Claims 6 and 16 are rejected under 35 U.S.C. 103(a) as being unpatentable over Pickering, in view of Sonmez et al., and further in view of Shriberg et al. (*Prosody-based Automatic Segmentation of Speech into Sentences and Topics*).

Pickering discloses tracking the mean (intermediate range) to recognize a slowly decreasing mean, signaling the end of a phrase (Fig. 4B, column 10, lines 6-13).

Neither Pickering nor Sonmez et al. discloses the at least one pitch parameter is a difference between the pitch information in the speech signal and baseline pitch information.

Shriberg et al. discloses determining a difference between pitch information in the speech signal and baseline information (the pitch range of a word relative to a baseline, page 135, 2<sup>nd</sup> column, 2<sup>nd</sup> paragraph, lines 1-5 and lines 11-16).

It would have been obvious to one of ordinary skill in the art at the time of invention to further modify the combination of Pickering and Sonmez et al. to determine a difference between pitch information and baseline information since the baseline is the most useful pitch parameter out of baselines, topline, and intermediate range measures, as taught by Shriberg et al. (page 135, 1<sup>st</sup> column, 2<sup>nd</sup> paragraph, lines 8-16).

9. Claims 7-9 and 17-19 are rejected under 35 U.S.C. 103(a) as being unpatentable over Pickering.

In regard to claims 7 and 17, Pickering discloses generating a posterior probability regarding the at least one speech endpoint.

The prosodic features are used to detect when the speaker has effectively finished talking (Fig. 3, step 560, column 9, lines 2-6 and lines 50-54). The test of step 560 indicates how likely the caller is to have finished (column 11, lines 2-9).

Official notice is taken that this likelihood would suggest to one of ordinary skill in the art at the time of invention using a posterior probability, since the well known likelihood function is a posterior probability.

In regard to claims 8 and 18, Pickering discloses the likelihood of a plurality of speaker states, including that a speaker has completed an utterance (finished speaking), that the speaker is pausing due to hesitation (the speaker will continue), and that the speaker is talking fluently (the speaker is in trouble and losing coherence, which would indicate that the speaker is not speaking fluently).

Pickering discloses the prosodic test at step 560 checks the pitch pattern for a long decline in pitch value at the end of an input, indicating the speaker is finished (column 10, lines 21-23); a final fall of short duration, which indicates the speaker is going to continue (column 10, lines 3-4); or a final rise with an excessively long duration, which indicates the speaker is in trouble and losing coherency (column 10, lines 4-5 and lines 28-29).

Thus, the examiner takes official notice that this would suggest to one of ordinary skill in the art at the time of invention to obtain the posterior probabilities that a that a

speaker has completed an utterance, that the speaker is pausing due to hesitation, and that the speaker is talking fluently, since the well known likelihood function is a posterior probability.

In regard to claims 9 and 19, Pickering's disclosed prosodic test at step 560 is based on a likelihood that the speaker is finished speaking (column 11, lines 2-9). If the speaker is not finished, steps 520-560 are repeated (see Fig. 3), which would suggest to one of ordinary skill in the art at the time of invention to update the posterior probability at step 560 as the speech signal is processed.

10. Claims 10 and 20 are rejected under 35 U.S.C. 103(a) as being unpatentable over Pickering, in view of Shin et al. (*Speech/Non-Speech Classification Using Multiple Features for Robust Endpoint Detection*).

Pickering discloses that after the prosodic test for an endpoint at step 560, further action is taken at step 570.

Pickering does not disclose that the further step is a speech recognition routine for processing the speech signal using the at least one speech endpoint.

Shin et al. discloses that the inaccurate detection of endpoints is a major cause of errors in speech recognition systems (page 1399, 1<sup>st</sup> column, section 1, 2<sup>nd</sup> paragraph, lines 1-3).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Pickering to perform speech recognition at step 570 using the speech endpoint, since increased endpoint detection accuracy increases the speech recognition performance, as taught by Shin et al. (page 1401, 1<sup>st</sup> column, section 4, 6<sup>th</sup> paragraph, lines 1-2 and page 1402, 1<sup>st</sup> column, 2<sup>nd</sup> paragraph, lines 7-10).

### ***Conclusion***

11. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. Lennig (U.S. Patent 6,873,953) disclose a method for prosody based endpoint detection.

12. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

13. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Brian L Albertalli whose telephone number is (571) 272-


Art Unit: 2655

7616. The examiner can normally be reached on Mon - Fri, 8:00 AM - 5:30 PM, every second Fri off.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Wayne Young can be reached on (571) 272-7582. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

BLA 5/23/05



DAVID L. OMETZ  
PRIMARY EXAMINER